

revised.SUB.SEQUENCE.LST.0120
SUBSTITUTE SEQUENCE LISTING



(1) GENERAL INFORMATION:

- (i) APPLICANT: PULST, STEFAN M.
- (ii) TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
ATAXIA-2 AND PRODUCTS RELATED THERETO
- (iii) NUMBER OF SEQUENCES: 19
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
 - (B) STREET: 119 NORTH FOURTH STREET, SUITE 203
 - (C) CITY: MINNEAPOLIS
 - (D) STATE: MINNESOTA
 - (E) COUNTRY: 55401
 - (F) ZIP: 55401
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/981,998
 - (B) FILING DATE: 11-MAY-1998
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: WO 97/42314
 - (B) FILING DATE: 08-MAY-1997
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/727,084
 - (B) FILING DATE: 08-OCT-1996
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/022,207
 - (B) FILING DATE: 19-JUL-1996
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/017,388
 - (B) FILING DATE: 08-MAY-1996
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: MUETING, ANN M.
 - (B) REGISTRATION NUMBER: 33,977
 - (C) REFERENCE/DOCKET NUMBER: 232.00010120
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 612-305-1217
 - (B) TELEFAX: 612-305-1228

RECEIVED

SEP 20 2001

TECH CENTER 1600/2900

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 516 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both

(D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

TTGGTAGCAA CGGAAACGGC GGC GGCGCGCT TTCGGCCCCG CTCCTTGGTC      60
TCGGCGGGCC TCCCCGCCCC TTCGTCGTCG TCCTTCTCCC CCTCGCCAGC CCGGGCGCCC      120
CTCCGGCCGC GCCAACCCGC GCCTCCCCGC TCGGCGCCCC TGCGTCCCCG CCGCGTTCCG      180
GCGTCTCCTT GGCGCGCCCG GCTCCCGGCT GTCCCCGCCC GGCGTGCGAG CCGGTGTATG      240
GGCCCCTCAC CATGTCGCTG AAGCCCCAGC AGCAGCAGCA GCAGCAGCAG CAACAGCAGC      300
AGCAGCAACA GCAGCAGCAG CAGCAGCAGC AGCCGCCGCC CGCGGCTGCC AATGTCCGCA      360
AGCCCGGCGG CAGCGGCCTT CTAGCGTCGC CCGCCGCCGC GCCTTCGCCG TCCTCGTCCT      420
CGGTCTCCTC GTCCTCGGCC ACGGCTCCCT CCTCGGTGGT CGCGGCGACC TCCGGCGGCG      480
GGAGGCCCCG CCTGGGCAGG TGGGTGTCGG CACCCC                                516

```

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4481 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: both
 (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 163..4101

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

ACCCCCGAGA AAGCAACCCA GCGCGCCGCC CGCTCCTCAC GTGTCCCTCC CGGCCCCGGG      60
GCCACCTCAC GTTCTGCTTC CGTCTGACCC CTCCGACTTC CGGTAAAGAG TCCCTATCCG      120
CACCTCCGCT CCCACCCGGC GCCTCGGCGC GCCCGCCCTC CG ATG CGC TCA GCG      174
                                     Met Arg Ser Ala
                                     1

GCC GCA GCT CCT CGG AGT CCC GCG GTG GCC ACC GAG TCT CGC CGC TTC      222
Ala Ala Ala Pro Arg Ser Pro Ala Val Ala Thr Glu Ser Arg Arg Phe
  5                10                15                20

GCC GCA GCC AGG TGG CCC GGG TGG CGC TCG CTC CAG CGG CCG GCG CGG      270
Ala Ala Ala Arg Trp Pro Gly Trp Arg Ser Leu Gln Arg Pro Ala Arg
                25                30                35

CGG AGC GGG CGG GGC GGC GGT GGC GCG GCC CCG GGA CCG TAT CCC TCC      318

```

revised.SUB.SEQUENCE.LST.0120

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|------|
| Arg | Ser | Gly | Arg | Gly | Gly | Gly | Gly | Ala | Ala | Pro | Gly | Pro | Tyr | Pro | Ser | | |
| | | | 40 | | | | | 45 | | | | | 50 | | | | |
| GCC | GCC | CCT | CCC | CCG | CCC | GGC | CCC | GGC | CCC | CCT | CCC | TCC | CGG | CAG | AGC | | 366 |
| Ala | Ala | Pro | Pro | Pro | Pro | Gly | Pro | Gly | Pro | Pro | Pro | Ser | Arg | Gln | Ser | | |
| | | 55 | | | | 60 | | | | | | 65 | | | | | |
| TCG | CCT | CCC | TCC | GCC | TCA | GAC | TGT | TTT | GGT | AGC | AAC | GGC | AAC | GGC | GGC | | 414 |
| Ser | Pro | Pro | Ser | Ala | Ser | Asp | Cys | Phe | Gly | Ser | Asn | Gly | Asn | Gly | Gly | | |
| | 70 | | | | | 75 | | | | | 80 | | | | | | |
| GGC | GCG | TTT | CGG | CCC | GGC | TCC | CGG | CGG | CTC | CTT | GGT | CTC | GGC | GGG | CCT | | 462 |
| Gly | Ala | Phe | Arg | Pro | Gly | Ser | Arg | Arg | Leu | Leu | Gly | Leu | Gly | Gly | Pro | | |
| | 85 | | | | 90 | | | | | 95 | | | | | 100 | | |
| CCC | CGC | CCC | TTC | GTC | GTC | GTC | CTT | CTC | CCC | CTC | GCC | AGC | CCG | GGC | GCC | | 510 |
| Pro | Arg | Pro | Phe | Val | Val | Val | Leu | Leu | Pro | Leu | Ala | Ser | Pro | Gly | Ala | | |
| | | | | 105 | | | | | 110 | | | | | 115 | | | |
| CCT | CCG | GCC | GCG | CCA | ACC | CGC | GCC | TCC | CCG | CTC | GGC | GCC | CGT | GCG | TCC | | 558 |
| Pro | Pro | Ala | Ala | Pro | Thr | Arg | Ala | Ser | Pro | Leu | Gly | Ala | Arg | Ala | Ser | | |
| | | | 120 | | | | | 125 | | | | | 130 | | | | |
| CCG | CCG | CGT | TCC | GGC | GTC | TCC | TTG | GCG | CGC | CCG | GCT | CCC | GGC | TGT | CCC | | 606 |
| Pro | Pro | Arg | Ser | Gly | Val | Ser | Leu | Ala | Arg | Pro | Ala | Pro | Gly | Cys | Pro | | |
| | | 135 | | | | | 140 | | | | | 145 | | | | | |
| CGC | CCG | GCG | TGC | GAG | CCG | GTG | TAT | GGG | CCC | CTC | ACC | ATG | TCG | CTG | AAG | | 654 |
| Arg | Pro | Ala | Cys | Glu | Pro | Val | Tyr | Gly | Pro | Leu | Thr | Met | Ser | Leu | Lys | | |
| | 150 | | | | | 155 | | | | | 160 | | | | | | |
| CCC | CAG | CAG | CAG | CAG | CAG | CAG | CAG | CAG | CAA | CAG | CAG | CAG | CAG | CAA | CAG | | 702 |
| Pro | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | | |
| | 165 | | | | 170 | | | | 175 | | | | | | 180 | | |
| CAG | CAG | CAG | CAG | CAG | CAG | CAG | CCG | CCG | CCC | GCG | GCT | GCC | AAT | GTC | CGC | | 750 |
| Gln | Gln | Gln | Gln | Gln | Gln | Gln | Pro | Pro | Pro | Ala | Ala | Ala | Asn | Val | Arg | | |
| | | | | 185 | | | | | 190 | | | | | 195 | | | |
| AAG | CCC | GGC | GGC | AGC | GGC | CTT | CTA | GCG | TCG | CCC | GCC | GCC | GCG | CCT | TCG | | 798 |
| Lys | Pro | Gly | Gly | Ser | Gly | Leu | Leu | Ala | Ser | Pro | Ala | Ala | Ala | Pro | Ser | | |
| | | | 200 | | | | | 205 | | | | | 210 | | | | |
| CCG | TCC | TCG | TCC | TCG | GTC | TCC | TCG | TCC | TCG | GCC | ACG | GCT | CCC | TCC | TCG | | 846 |
| Pro | Ser | Ser | Ser | Ser | Val | Ser | Ser | Ser | Ser | Ala | Thr | Ala | Pro | Ser | Ser | | |
| | | 215 | | | | 220 | | | | | | 225 | | | | | |
| GTG | GTC | GCG | GCG | ACC | TCC | GGC | GGC | GGG | AGG | CCC | GGC | CTG | GGC | AGA | GGT | | 894 |
| Val | Val | Ala | Ala | Thr | Ser | Gly | Gly | Gly | Arg | Pro | Gly | Leu | Gly | Arg | Gly | | |
| | 230 | | | | | 235 | | | | | 240 | | | | | | |
| CGA | AAC | AGT | AAC | AAA | GGA | CTG | CCT | CAG | TCT | ACG | ATT | TCT | TTT | GAT | GGA | | 942 |
| Arg | Asn | Ser | Asn | Lys | Gly | Leu | Pro | Gln | Ser | Thr | Ile | Ser | Phe | Asp | Gly | | |
| | 245 | | | | 250 | | | | | 255 | | | | | 260 | | |
| ATC | TAT | GCA | AAT | ATG | AGG | ATG | GTT | CAT | ATA | CTT | ACA | TCA | GTT | GTT | GGC | | 990 |
| Ile | Tyr | Ala | Asn | Met | Arg | Met | Val | His | Ile | Leu | Thr | Ser | Val | Val | Gly | | |
| | | | | 265 | | | | | 270 | | | | | 275 | | | |
| TCC | AAA | TGT | GAA | GTA | CAA | GTG | AAA | AAT | GGA | GGT | ATA | TAT | GAA | GGA | GTT | | 1038 |
| Ser | Lys | Cys | Glu | Val | Gln | Val | Lys | Asn | Gly | Gly | Ile | Tyr | Glu | Gly | Val | | |

revised.SUB.SEQUENCE.LST.0120

| 280 | 285 | 290 | |
|---|-----|-----|------|
| TTT AAA ACT TAC AGT CCG AAG TGT GAT TTG GTA CTT GAT GCC GCA CAT Phe Lys Thr Tyr Ser Pro Lys Cys Asp Leu Val Leu Asp Ala Ala His 295 300 305 | | | 1086 |
| GAG AAA AGT ACA GAA TCC AGT TCG GGG CCG AAA CGT GAA GAA ATA ATG Glu Lys Ser Thr Glu Ser Ser Ser Gly Pro Lys Arg Glu Glu Ile Met 310 315 320 | | | 1134 |
| GAG AGT ATT TTG TTC AAA TGT TCA GAC TTT GTT GTG GTA CAG TTT AAA Glu Ser Ile Leu Phe Lys Cys Ser Asp Phe Val Val Val Gln Phe Lys 325 330 335 | | | 1182 |
| GAT ATG GAC TCC AGT TAT GCA AAA AGA GAT GCT TTT ACT GAC TCT GCT Asp Met Asp Ser Ser Tyr Ala Lys Arg Asp Ala Phe Thr Asp Ser Ala 345 350 355 | | | 1230 |
| ATC AGT GCT AAA GTG AAT GGC GAA CAC AAA GAG AAG GAC CTG GAG CCC Ile Ser Ala Lys Val Asn Gly Glu His Lys Glu Lys Asp Leu Glu Pro 360 365 370 | | | 1278 |
| TGG GAT GCA GGT GAA CTC ACA GCC AAT GAG GAA CTT GAG GCT TTG GAA Trp Asp Ala Gly Glu Leu Thr Ala Asn Glu Glu Leu Glu Ala Leu Glu 375 380 385 | | | 1326 |
| AAT GAC GTA TCT AAT GGA TGG GAT CCC AAT GAT ATG TTT CGA TAT AAT Asn Asp Val Ser Asn Gly Trp Asp Pro Asn Asp Met Phe Arg Tyr Asn 390 395 400 | | | 1374 |
| GAA GAA AAT TAT GGT GTA GTG TCT ACG TAT GAT AGC AGT TTA TCT TCG Glu Glu Asn Tyr Gly Val Val Ser Thr Tyr Asp Ser Ser Leu Ser Ser 405 410 415 420 | | | 1422 |
| TAT ACA GTG CCC TTA GAA AGA GAT AAC TCA GAA GAA TTT TTA AAA CGG Tyr Thr Val Pro Leu Glu Arg Asp Asn Ser Glu Glu Phe Leu Lys Arg 425 430 435 | | | 1470 |
| GAA GCA AGG GCA AAC CAG TTA GCA GAA GAA ATT GAG TCA AGT GCC CAG Glu Ala Arg Ala Asn Gln Leu Ala Glu Glu Ile Glu Ser Ser Ala Gln 440 445 450 | | | 1518 |
| TAC AAA GCT CGA GTG GCC CTG GAA AAT GAT GAT AGG AGT GAG GAA GAA Tyr Lys Ala Arg Val Ala Leu Glu Asn Asp Asp Arg Ser Glu Glu Glu 455 460 465 | | | 1566 |
| AAA TAC ACA GCA GTT CAG AGA AAT TCC AGT GAA CGT GAG GGG CAC AGC Lys Tyr Thr Ala Val Gln Arg Asn Ser Ser Glu Arg Glu Gly His Ser 470 475 480 | | | 1614 |
| ATA AAC ACT AGG GAA AAT AAA TAT ATT CCT CCT GGA CAA AGA AAT AGA Ile Asn Thr Arg Glu Asn Lys Tyr Ile Pro Pro Gly Gln Arg Asn Arg 485 490 495 500 | | | 1662 |
| GAA GTC ATA TCC TGG GGA AGT GGG AGA CAG AAT TCA CCG CGT ATG GGC Glu Val Ile Ser Trp Gly Ser Gly Arg Gln Asn Ser Pro Arg Met Gly 505 510 515 | | | 1710 |
| CAG CCT GGA TCG GGC TCC ATG CCA TCA AGA TCC ACT TCT CAC ACT TCA Gln Pro Gly Ser Gly Ser Met Pro Ser Arg Ser Thr Ser His Thr Ser 520 525 530 | | | 1758 |

revised.SUB.SEQUENCE.LST.0120

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| GAT | TTC | AAC | CCG | AAT | TCT | GGT | TCA | GAC | CAA | AGA | GTA | GTT | AAT | GGA | GGT | 1806 |
| Asp | Phe | Asn | Pro | Asn | Ser | Gly | Ser | Asp | Gln | Arg | Val | Val | Asn | Gly | Gly | |
| | | 535 | | | | | 540 | | | | | 545 | | | | |
| GTT | CCC | TGG | CCA | TCG | CCT | TGC | CCA | TCT | CCT | TCC | TCT | CGC | CCA | CCT | TCT | 1854 |
| Val | Pro | Trp | Pro | Ser | Pro | Cys | Pro | Ser | Pro | Ser | Ser | Arg | Pro | Pro | Ser | |
| | 550 | | | | | 555 | | | | | 560 | | | | | |
| CGC | TAC | CAG | TCA | GGT | CCC | AAC | TCT | CTT | CCA | CCT | CGG | GCA | GCC | ACC | CCT | 1902 |
| Arg | Tyr | Gln | Ser | Gly | Pro | Asn | Ser | Leu | Pro | Pro | Arg | Ala | Ala | Thr | Pro | |
| 565 | | | | | 570 | | | | | 575 | | | | | 580 | |
| ACA | CGG | CCG | CCC | TCC | AGG | CCC | CCC | TCG | CGG | CCA | TCC | AGA | CCC | CCG | TCT | 1950 |
| Thr | Arg | Pro | Pro | Ser | Arg | Pro | Pro | Ser | Arg | Pro | Ser | Arg | Pro | Pro | Ser | |
| | | | | 585 | | | | | 590 | | | | | 595 | | |
| CAC | CCC | TCT | GCT | CAT | GGT | TCT | CCA | GCT | CCT | GTC | TCT | ACT | ATG | CCT | AAA | 1998 |
| His | Pro | Ser | Ala | His | Gly | Ser | Pro | Ala | Pro | Val | Ser | Thr | Met | Pro | Lys | |
| | | | 600 | | | | | 605 | | | | | 610 | | | |
| CGC | ATG | TCT | TCA | GAA | GGG | CCT | CCA | AGG | ATG | TCC | CCA | AAG | GCC | CAG | CGA | 2046 |
| Arg | Met | Ser | Ser | Glu | Gly | Pro | Pro | Arg | Met | Ser | Pro | Lys | Ala | Gln | Arg | |
| | | 615 | | | | | 620 | | | | | 625 | | | | |
| CAT | CCT | CGA | AAT | CAC | AGA | GTT | TCT | GCT | GGG | AGG | GGT | TCC | ATA | TCC | AGT | 2094 |
| His | Pro | Arg | Asn | His | Arg | Val | Ser | Ala | Gly | Arg | Gly | Ser | Ile | Ser | Ser | |
| | 630 | | | | | 635 | | | | | 640 | | | | | |
| GGC | CTA | GAA | TTT | GTA | TCC | CAC | AAC | CCA | CCC | AGT | GAA | GCA | GCT | ACT | CCT | 2142 |
| Gly | Leu | Glu | Phe | Val | Ser | His | Asn | Pro | Pro | Ser | Glu | Ala | Ala | Thr | Pro | |
| 645 | | | | | 650 | | | | | 655 | | | | | 660 | |
| CCA | GTA | GCA | AGG | ACC | AGT | CCC | TCG | GGG | GGA | ACG | TGG | TCA | TCA | GTG | GTC | 2190 |
| Pro | Val | Ala | Arg | Thr | Ser | Pro | Ser | Gly | Gly | Thr | Trp | Ser | Ser | Val | Val | |
| | | | | 665 | | | | | 670 | | | | | 675 | | |
| AGT | GGG | GTT | CCA | AGA | TTA | TCC | CCT | AAA | ACT | CAT | AGA | CCC | AGG | TCT | CCC | 2238 |
| Ser | Gly | Val | Pro | Arg | Leu | Ser | Pro | Lys | Thr | His | Arg | Pro | Arg | Ser | Pro | |
| | | | 680 | | | | | 685 | | | | | 690 | | | |
| AGA | CAG | AAC | AGT | ATT | GGA | AAT | ACC | CCC | AGT | GGG | CCA | GTT | CTT | GCT | TCT | 2286 |
| Arg | Gln | Asn | Ser | Ile | Gly | Asn | Thr | Pro | Ser | Gly | Pro | Val | Leu | Ala | Ser | |
| | | 695 | | | | | 700 | | | | | 705 | | | | |
| CCC | CAA | GCT | GGT | ATT | ATT | CCA | ACT | GAA | GCT | GTT | GCC | ATG | CCT | ATT | CCA | 2334 |
| Pro | Gln | Ala | Gly | Ile | Ile | Pro | Thr | Glu | Ala | Val | Ala | Met | Pro | Ile | Pro | |
| | 710 | | | | | 715 | | | | | 720 | | | | | |
| GCT | GCA | TCT | CCT | ACG | CCT | GCT | AGT | CCT | GCA | TCG | AAC | AGA | GCT | GTT | ACC | 2382 |
| Ala | Ala | Ser | Pro | Thr | Pro | Ala | Ser | Pro | Ala | Ser | Asn | Arg | Ala | Val | Thr | |
| 725 | | | | | 730 | | | | | 735 | | | | | 740 | |
| CCT | TCT | AGT | GAG | GCT | AAA | GAT | TCC | AGG | CTT | CAA | GAT | CAG | AGG | CAG | AAC | 2430 |
| Pro | Ser | Ser | Glu | Ala | Lys | Asp | Ser | Arg | Leu | Gln | Asp | Gln | Arg | Gln | Asn | |
| | | | | 745 | | | | | 750 | | | | | 755 | | |
| TCT | CCT | GCA | GGG | AAT | AAA | GAA | AAT | ATT | AAA | CCC | AAT | GAA | ACA | TCA | CCT | 2478 |
| Ser | Pro | Ala | Gly | Asn | Lys | Glu | Asn | Ile | Lys | Pro | Asn | Glu | Thr | Ser | Pro | |
| | | | 760 | | | | | 765 | | | | | 770 | | | |

revised.SUB.SEQUENCE.LST.0120

| | | | | | | | | | | | | | | | | |
|-------------------|-------------------|-------------------|--------------------|-------------------|-------------------|-------------------|-------------------|--------------------|-------------------|-------------------|-------------------|-------------------|--------------------|-------------------|-------------------|------|
| AGC Ser | TTC Phe | TCA Ser 775 | AAA Lys | GCT Ala | GAA Glu | AAC Asn | AAA Lys 780 | GGT Gly | ATA Ile | TCA Ser | CCA Pro | GTT Val 785 | GTT Val | TCT Ser | GAA Glu | 2526 |
| CAT His | AGA Arg 790 | AAA Lys | CAG Gln | ATT Ile | GAT Asp | GAT Asp 795 | TTA Leu | AAG Lys | AAA Lys | TTT Phe | AAG Lys 800 | AAT Asn | GAT Asp | TTT Phe | AGG Arg | 2574 |
| TTA Leu 805 | CAG Gln | CCA Pro | AGT Ser | TCT Ser | ACT Thr 810 | TCT Ser | GAA Glu | TCT Ser | ATG Met | GAT Asp 815 | CAA Gln | CTA Leu | CTA Leu | AAC Asn | AAA Lys 820 | 2622 |
| AAT Asn | AGA Arg | GAG Glu | GGA Gly | GAA Glu 825 | AAA Lys | TCA Ser | AGA Arg | GAT Asp | TTG Leu 830 | ATC Ile | AAA Lys | GAC Asp | AAA Lys | ATT Ile 835 | GAA Glu | 2670 |
| CCA Pro | AGT Ser | GCT Ala | AAG Lys 840 | GAT Asp | TCT Ser | TTC Phe | ATT Ile | GAA Glu 845 | AAT Asn | AGC Ser | AGC Ser | AGC Ser | AAC Asn 850 | TGT Cys | ACC Thr | 2718 |
| AGT Ser | GGC Gly | AGC Ser 855 | AGC Ser | AAG Lys | CCG Pro | AAT Asn | AGC Ser 860 | CCC Pro | AGC Ser | ATT Ile | TCC Ser | CCT Pro 865 | TCA Ser | ATA Ile | CTT Leu | 2766 |
| AGT Ser | AAC Asn 870 | ACG Thr | GAG Glu | CAC His | AAG Lys | AGG Arg 875 | GGA Gly | CCT Pro | GAG Glu | GTC Val | ACT Thr 880 | TCC Ser | CAA Gln | GGG Gly | GTT Val | 2814 |
| CAG Gln 885 | ACT Thr | TCC Ser | AGC Ser | CCA Pro | GCA Ala 890 | TGT Cys | AAA Lys | CAA Gln | GAG Glu | AAA Lys 895 | GAC Asp | GAT Asp | AAG Lys | GAA Glu | GAG Glu 900 | 2862 |
| AAG Lys | AAA Lys | GAC Asp | GCA Ala | GCT Ala 905 | GAG Glu | CAA Gln | GTT Val | AGG Arg | AAA Lys 910 | TCA Ser | ACA Thr | TTG Leu | AAT Asn | CCC Pro 915 | AAT Asn | 2910 |
| GCA Ala | AAG Lys | GAG Glu | TTC Phe 920 | AAC Asn | CCA Pro | CGT Arg | TCC Ser | TTC Phe 925 | TCT Ser | CAG Gln | CCA Pro | AAG Lys | CCT Pro 930 | TCT Ser | ACT Thr | 2958 |
| ACC Thr | CCA Pro | ACT Thr 935 | TCA Ser | CCT Pro | CGG Arg | CCT Pro | CAA Gln 940 | GCA Ala | CAA Gln | CCT Pro | AGC Ser | CCA Pro 945 | TCT Ser | ATG Met | GTG Val | 3006 |
| GGT Gly | CAT His 950 | CAA Gln | CAG Gln | CCA Pro | ACT Thr | CCA Pro 955 | GTT Val | TAT Tyr | ACT Thr | CAG Gln | CCT Pro 960 | GTT Val | TGT Cys | TTT Phe | GCA Ala | 3054 |
| CCA Pro 965 | AAT Asn | ATG Met | ATG Met | TAT Tyr | CCA Pro 970 | GTC Val | CCA Pro | GTG Val | AGC Ser | CCA Pro 975 | GGC Gly | GTG Val | CAA Gln | CCT Pro | TTA Leu 980 | 3102 |
| TAC Tyr | CCA Pro | ATA Ile | CCT Pro | ATG Met 985 | ACG Thr | CCC Pro | ATG Met | CCA Pro | GTG Val 990 | AAT Asn | CAA Gln | GCC Ala | AAG Lys | ACA Thr 995 | TAT Tyr | 3150 |
| AGA Arg | GCA Ala | GTA Val | CCA Pro 1000 | AAT Asn | ATG Met | CCC Pro | CAA Gln | CAG Gln 1005 | CGG Arg | CAA Gln | GAC Asp | CAG Gln | CAT His 1010 | CAT His | CAG Gln | 3198 |
| AGT | GCC | ATG | ATG | CAC | CCA | GCG | TCA | GCA | GCG | GGC | CCA | CCG | ATT | GCA | GCC | 3246 |

revised.SUB.SEQUENCE.LST.0120

| | | |
|---|-------------------------------------|------|
| Ser Ala Met Met His Pro Ala | Ser Ala Ala Gly Pro Pro Ile Ala Ala | |
| 1015 | 1020 1025 | |
| ACC CCA CCA GCT TAC TCC ACG CAA TAT GTT GCC TAC AGT CCT CAG CAG | | 3294 |
| Thr Pro Pro Ala Tyr Ser Thr Gln Tyr Val Ala Tyr Ser Pro Gln Gln | 1030 1035 1040 | |
| TTC CCA AAT CAG CCC CTT GTT CAG CAT GTG CCA CAT TAT CAG TCT CAG | | 3342 |
| Phe Pro Asn Gln Pro Leu Val Gln His Val Pro His Tyr Gln Ser Gln | 1045 1050 1055 1060 | |
| CAT CCT CAT GTC TAT AGT CCT GTA ATA CAG GGT AAT GCT AGA ATG ATG | | 3390 |
| His Pro His Val Tyr Ser Pro Val Ile Gln Gly Asn Ala Arg Met Met | 1065 1070 1075 | |
| GCA CCA CCA ACA CAC GCC CAG CCT GGT TTA GTA TCT TCT TCA GCA ACT | | 3438 |
| Ala Pro Pro Thr His Ala Gln Pro Gly Leu Val Ser Ser Ser Ala Thr | 1080 1085 1090 | |
| CAG TAC GGG GCT CAT GAG CAG ACG CAT GCG ATG TAT GCA TGT CCC AAA | | 3486 |
| Gln Tyr Gly Ala His Glu Gln Thr His Ala Met Tyr Ala Cys Pro Lys | 1095 1100 1105 | |
| TTA CCA TAC AAC AAG GAG ACA AGC CCT TCT TTC TAC TTT GCC ATT TCC | | 3534 |
| Leu Pro Tyr Asn Lys Glu Thr Ser Pro Ser Phe Tyr Phe Ala Ile Ser | 1110 1115 1120 | |
| ACG GGC TCC CTT GCT CAG CAG TAT GCG CAC CCT AAC GCT ACC CTG CAC | | 3582 |
| Thr Gly Ser Leu Ala Gln Gln Tyr Ala His Pro Asn Ala Thr Leu His | 1125 1130 1135 1140 | |
| CCA CAT ACT CCA CAC CCT CAG CCT TCA GCT ACC CCC ACT GGA CAG CAG | | 3630 |
| Pro His Thr Pro His Pro Gln Pro Ser Ala Thr Pro Thr Gly Gln Gln | 1145 1150 1155 | |
| CAA AGC CAA CAT GGT GGA AGT CAT CCT GCA CCC AGT CCT GTT CAG CAC | | 3678 |
| Gln Ser Gln His Gly Gly Ser His Pro Ala Pro Ser Pro Val Gln His | 1160 1165 1170 | |
| CAT CAG CAC CAG GCC GCC CAG GCT CTC CAT CTG GCC AGT CCA CAG CAG | | 3726 |
| His Gln His Gln Ala Ala Gln Ala Leu His Leu Ala Ser Pro Gln Gln | 1175 1180 1185 | |
| CAG TCA GCC ATT TAC CAC GCG GGG CTT GCG CCA ACT CCA CCC TCC ATG | | 3774 |
| Gln Ser Ala Ile Tyr His Ala Gly Leu Ala Pro Thr Pro Pro Ser Met | 1190 1195 1200 | |
| ACA CCT GCC TCC AAC ACG CAG TCG CCA CAG AAT AGT TTC CCA GCA GCA | | 3822 |
| Thr Pro Ala Ser Asn Thr Gln Ser Pro Gln Asn Ser Phe Pro Ala Ala | 1205 1210 1215 1220 | |
| CAA CAG ACT GTC TTT ACG ATC CAT CCT TCT CAC GTT CAG CCG GCG TAT | | 3870 |
| Gln Gln Thr Val Phe Thr Ile His Pro Ser His Val Gln Pro Ala Tyr | 1225 1230 1235 | |
| ACC AAC CCA CCC CAC ATG GCC CAC GTA CCT CAG GCT CAT GTA CAG TCA | | 3918 |
| Thr Asn Pro Pro His Met Ala His Val Pro Gln Ala His Val Gln Ser | 1240 1245 1250 | |
| GGA ATG GTT CCT TCT CAT CCA ACT GCC CAT GCG CCA ATG ATG CTA ATG | | 3966 |
| Gly Met Val Pro Ser His Pro Thr Ala His Ala Pro Met Met Leu Met | | |

revised.SUB.SEQUENCE.LST.0120
1260 1265

| | | |
|------|--|------|
| 1255 | ACG ACA CAG CCA CCC GGC GGT CCC CAG GCC GCC CTC GCT CAA AGT GCA | 4014 |
| | Thr Thr Gln Pro Pro Gly Gly Pro Gln Ala Ala Leu Ala Gln Ser Ala | |
| | 1270 1275 1280 | |
| | CTA CAG CCC ATT CCA GTC TCG ACA ACA GCG CAT TTC CCC TAT ATG ACG | 4062 |
| | Leu Gln Pro Ile Pro Val Ser Thr Thr Ala His Phe Pro Tyr Met Thr | |
| | 1285 1290 1295 1300 | |
| | CAC CCT TCA GTA CAA GCC CAC CAC CAA CAG CAG TTG TAA GGCTGCCCTG | 4111 |
| | His Pro Ser Val Gln Ala His His Gln Gln Gln Leu * | |
| | 1305 1310 | |
| | GAGGAACCGA AAGGCCAAAT TCCCTCCTCC CTTCTACTGC TTCTACCAAC TGGGAAGCACA | 4171 |
| | GAAAACTAGA ATTTCAATTA TTTTGTTTT AAAATATATA TGTTGATTTC TTGTAACATC | 4231 |
| | CAATAGGAAT GCTAACAGTT CACTTGCACT GGAAGATACT TGGACCGAGT AGAGGCATTT | 4291 |
| | AGGAACTTGG GGGCTATTCC ATAATTCCAT ATGCTGTTTC AGAGTCCCGC AGGTACCCCA | 4351 |
| | GCTCTGCTTG CCGAACTGG AAGTTATTTA TTTTITAATA ACCCTTGAAA GTCATGAACA | 4411 |
| | CATCAGCTAG CAAAAGAAGT AACAAGAGTG ATTCTTGCTG CTATTACTGC TAAAAA | 4471 |
| | AAAAA | 4481 |

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1312 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Arg Ser Ala Ala Ala Ala Pro Arg Ser Pro Ala Val Ala Thr Glu
1 5 10 15

Ser Arg Arg Phe Ala Ala Ala Arg Trp Pro Gly Trp Arg Ser Leu Gln
20 25 30

Arg Pro Ala Arg Arg Ser Gly Arg Gly Gly Gly Gly Ala Ala Pro Gly
35 40 45

Pro Tyr Pro Ser Ala Ala Pro Pro Pro Gly Pro Gly Pro Pro Pro
50 55 60

Ser Arg Gln Ser Ser Pro Pro Ser Ala Ser Asp Cys Phe Gly Ser Asn
65 70 75 80

Gly Asn Gly Gly Gly Ala Phe Arg Pro Gly Ser Arg Arg Leu Leu Gly
85 90 95

Leu Gly Gly Pro Pro Arg Pro Phe Val Val Val Leu Leu Pro Leu Ala
100 105 110

Ser Pro Gly Ala Pro Pro Ala Ala Pro Thr Arg Ala Ser Pro Leu Gly
115 120 125

revised.SUB.SEQUENCE.LST.0120

Ala Arg Ala Ser Pro Pro Arg Ser Gly Val Ser Leu Ala Arg Pro Ala
130 135 140
Pro Gly Cys Pro Arg Pro Ala Cys Glu Pro Val Tyr Gly Pro Leu Thr
145 150 155 160
Met Ser Leu Lys Pro Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
165 170 175
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Pro Pro Pro Ala Ala
180 185 190
Ala Asn Val Arg Lys Pro Gly Gly Ser Gly Leu Leu Ala Ser Pro Ala
195 200 205
Ala Ala Pro Ser Pro Ser Ser Ser Ser Val Ser Ser Ser Ser Ala Thr
210 215 220
Ala Pro Ser Ser Val Val Ala Ala Thr Ser Gly Gly Gly Arg Pro Gly
225 230 235 240
Leu Gly Arg Gly Arg Asn Ser Asn Lys Gly Leu Pro Gln Ser Thr Ile
245 250 255
Ser Phe Asp Gly Ile Tyr Ala Asn Met Arg Met Val His Ile Leu Thr
260 265 270
Ser Val Val Gly Ser Lys Cys Glu Val Gln Val Lys Asn Gly Gly Ile
275 280 285
Tyr Glu Gly Val Phe Lys Thr Tyr Ser Pro Lys Cys Asp Leu Val Leu
290 295 300
Asp Ala Ala His Glu Lys Ser Thr Glu Ser Ser Ser Gly Pro Lys Arg
305 310 315 320
Glu Glu Ile Met Glu Ser Ile Leu Phe Lys Cys Ser Asp Phe Val Val
325 330 335
Val Gln Phe Lys Asp Met Asp Ser Ser Tyr Ala Lys Arg Asp Ala Phe
340 345 350
Thr Asp Ser Ala Ile Ser Ala Lys Val Asn Gly Glu His Lys Glu Lys
355 360 365
Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu Thr Ala Asn Glu Glu Leu
370 375 380
Glu Ala Leu Glu Asn Asp Val Ser Asn Gly Trp Asp Pro Asn Asp Met
385 390 395 400
Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val Val Ser Thr Tyr Asp Ser
405 410 415
Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Arg Asp Asn Ser Glu Glu
420 425 430
Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln Leu Ala Glu Glu Ile Glu
435 440 445
Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala Leu Glu Asn Asp Asp Arg
450 455 460

revised.SUB.SEQUENCE.LST.0120

Ser Glu Glu Glu Lys Tyr Thr Ala Val Gln Arg Asn Ser Ser Glu Arg
465 470 475 480

Glu Gly His Ser Ile Asn Thr Arg Glu Asn Lys Tyr Ile Pro Pro Gly
485 490 495

Gln Arg Asn Arg Glu Val Ile Ser Trp Gly Ser Gly Arg Gln Asn Ser
500 505 510

Pro Arg Met Gly Gln Pro Gly Ser Gly Ser Met Pro Ser Arg Ser Thr
515 520 525

Ser His Thr Ser Asp Phe Asn Pro Asn Ser Gly Ser Asp Gln Arg Val
530 535 540

Val Asn Gly Gly Val Pro Trp Pro Ser Pro Cys Pro Ser Pro Ser Ser
545 550 555 560

Arg Pro Pro Ser Arg Tyr Gln Ser Gly Pro Asn Ser Leu Pro Pro Arg
565 570 575

Ala Ala Thr Pro Thr Arg Pro Pro Ser Arg Pro Pro Ser Arg Pro Ser
580 585 590

Arg Pro Pro Ser His Pro Ser Ala His Gly Ser Pro Ala Pro Val Ser
595 600 605

Thr Met Pro Lys Arg Met Ser Ser Glu Gly Pro Pro Arg Met Ser Pro
610 615 620

Lys Ala Gln Arg His Pro Arg Asn His Arg Val Ser Ala Gly Arg Gly
625 630 635 640

Ser Ile Ser Ser Gly Leu Glu Phe Val Ser His Asn Pro Pro Ser Glu
645 650 655

Ala Ala Thr Pro Pro Val Ala Arg Thr Ser Pro Ser Gly Gly Thr Trp
660 665 670

Ser Ser Val Val Ser Gly Val Pro Arg Leu Ser Pro Lys Thr His Arg
675 680 685

Pro Arg Ser Pro Arg Gln Asn Ser Ile Gly Asn Thr Pro Ser Gly Pro
690 695 700

Val Leu Ala Ser Pro Gln Ala Gly Ile Ile Pro Thr Glu Ala Val Ala
705 710 715 720

Met Pro Ile Pro Ala Ala Ser Pro Thr Pro Ala Ser Pro Ala Ser Asn
725 730 735

Arg Ala Val Thr Pro Ser Ser Glu Ala Lys Asp Ser Arg Leu Gln Asp
740 745 750

Gln Arg Gln Asn Ser Pro Ala Gly Asn Lys Glu Asn Ile Lys Pro Asn
755 760 765

Glu Thr Ser Pro Ser Phe Ser Lys Ala Glu Asn Lys Gly Ile Ser Pro
770 775 780

Val Val Ser Glu His Arg Lys Gln Ile Asp Asp Leu Lys Lys Phe Lys

revised.SUB.SEQUENCE.LST.0120

785 790 795 800
 Asn Asp Phe Arg Leu Gln Pro Ser Ser Thr Ser Glu Ser Met Asp Gln
 805 810 815
 Leu Leu Asn Lys Asn Arg Glu Gly Glu Lys Ser Arg Asp Leu Ile Lys
 820 825 830
 Asp Lys Ile Glu Pro Ser Ala Lys Asp Ser Phe Ile Glu Asn Ser Ser
 835 840 845
 Ser Asn Cys Thr Ser Gly Ser Ser Lys Pro Asn Ser Pro Ser Ile Ser
 850 855 860
 Pro Ser Ile Leu Ser Asn Thr Glu His Lys Arg Gly Pro Glu Val Thr
 865 870 875 880
 Ser Gln Gly Val Gln Thr Ser Ser Pro Ala Cys Lys Gln Glu Lys Asp
 885 890 895
 Asp Lys Glu Glu Lys Lys Asp Ala Ala Glu Gln Val Arg Lys Ser Thr
 900 905 910
 Leu Asn Pro Asn Ala Lys Glu Phe Asn Pro Arg Ser Phe Ser Gln Pro
 915 920 925
 Lys Pro Ser Thr Thr Pro Thr Ser Pro Arg Pro Gln Ala Gln Pro Ser
 930 935 940
 Pro Ser Met Val Gly His Gln Gln Pro Thr Pro Val Tyr Thr Gln Pro
 945 950 955 960
 Val Cys Phe Ala Pro Asn Met Met Tyr Pro Val Pro Val Ser Pro Gly
 965 970 975
 Val Gln Pro Leu Tyr Pro Ile Pro Met Thr Pro Met Pro Val Asn Gln
 980 985 990
 Ala Lys Thr Tyr Arg Ala Val Pro Asn Met Pro Gln Gln Arg Gln Asp
 995 1000 1005
 Gln His His Gln Ser Ala Met Met His Pro Ala Ser Ala Ala Gly Pro
 1010 1015 1020
 Pro Ile Ala Ala Thr Pro Pro Ala Tyr Ser Thr Gln Tyr Val Ala Tyr
 1025 1030 1035 1040
 Ser Pro Gln Gln Phe Pro Asn Gln Pro Leu Val Gln His Val Pro His
 1045 1050 1055
 Tyr Gln Ser Gln His Pro His Val Tyr Ser Pro Val Ile Gln Gly Asn
 1060 1065 1070
 Ala Arg Met Met Ala Pro Pro Thr His Ala Gln Pro Gly Leu Val Ser
 1075 1080 1085
 Ser Ser Ala Thr Gln Tyr Gly Ala His Glu Gln Thr His Ala Met Tyr
 1090 1095 1100
 Ala Cys Pro Lys Leu Pro Tyr Asn Lys Glu Thr Ser Pro Ser Phe Tyr
 1105 1110 1115 1120

revised.SUB.SEQUENCE.LST.0120

Phe Ala Ile Ser Thr Gly Ser Leu Ala Gln Gln Tyr Ala His Pro Asn
1125 1130 1135

Ala Thr Leu His Pro His Thr Pro His Pro Gln Pro Ser Ala Thr Pro
1140 1145 1150

Thr Gly Gln Gln Gln Ser Gln His Gly Gly Ser His Pro Ala Pro Ser
1155 1160 1165

Pro Val Gln His His Gln His Gln Ala Ala Gln Ala Leu His Leu Ala
1170 1175 1180

Ser Pro Gln Gln Gln Ser Ala Ile Tyr His Ala Gly Leu Ala Pro Thr
1185 1190 1195 1200

Pro Pro Ser Met Thr Pro Ala Ser Asn Thr Gln Ser Pro Gln Asn Ser
1205 1210 1215

Phe Pro Ala Ala Gln Gln Thr Val Phe Thr Ile His Pro Ser His Val
1220 1225 1230

Gln Pro Ala Tyr Thr Asn Pro Pro His Met Ala His Val Pro Gln Ala
1235 1240 1245

His Val Gln Ser Gly Met Val Pro Ser His Pro Thr Ala His Ala Pro
1250 1255 1260

Met Met Leu Met Thr Thr Gln Pro Pro Gly Gly Pro Gln Ala Ala Leu
1265 1270 1275 1280

Ala Gln Ser Ala Leu Gln Pro Ile Pro Val Ser Thr Thr Ala His Phe
1285 1290 1295

Pro Tyr Met Thr His Pro Ser Val Gln Ala His His Gln Gln Gln Leu
1300 1305 1310

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3798 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 50..3457

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGCACGAGGT CCCC GCCCGG CGTGCGAGCC GGTGTATGGG CCGCTCACC ATG TCG 55
Met Ser
1

CTG AAG CCG CAG CCG CAG CCG CCC GCG CCC GCC ACT GGC CGC AAG CCC 103
Page 12

Leu Lys Pro Gln Pro Gln Pro Pro Ala Pro Ala Thr Gly Arg Lys Pro
5 10 15

GTG ACC TCG GCT TCC GTG GTG CCG GCC CCG GCC GCG CCG GTG GCG TCT
Val Thr Ser Ala Ser Val Val Pro Ala Pro Ala Ala Pro Val Ala Ser
35 40 45 50 199

AAC AGT AGC AAA GGA CTG CCT CAG CCT ACG ATT TCT TTT GAT GGA ATC 295
Asn Ser Ser Lys Gly Leu Pro Gln Pro Thr Ile Ser Phe Asp Gly Ile
70 75 80

AAA TGT GAA GTA CAA GTG AAA AAC GGA GGC ATA TAT GAA GGA GTT TTT 391
Lys Cys Glu Val Gln Val Lys Asn Gly Gly Ile Tyr Glu Gly Val Phe
100 105 110

AAA AGT ACA GAA TCC AGT TCG GGG CCA AAA CGT GAA GAA ATA ATG GAG 487
Lys Ser Thr Glu Ser Ser Ser Gly Pro Lys Arg Glu Glu Ile Met Glu
135 140 145

ACA GAC TCC AGT TAT GCA CGG AGA GAT GCT TTT ACT GAC TCT GCT CTC 583
Thr Asp Ser Ser Tyr Ala Arg Arg Asp Ala Phe Thr Asp Ser Ala Leu
165 170 175

GAT GCA GGG GAG CTC ACG GCC AGC GAG GAG CTG GAG CTG GAG AAT GAT
Asp Ala Gly Glu Leu Thr Ala Ser Glu Glu Leu Glu Leu Glu Asn Asp
195 200 205 210 679

AAT TAT GGT GTG GTG TCC ACA TAT GAT AGC AGT TTA TCT TCA TAT ACG 775
Asn Tyr Gly Val Val Ser Thr Gln Ser Ser Tyr Thr

Val Pro Leu Glu Arg Asp Asn Ser Glu Glu Phe Leu Lys Arg Glu Ala
Page 13

revised.SUB.SEQUENCE.LST.0120
250 255

| | | |
|-----|---|------|
| 245 | AGG GCA AAC CAG TTA GCA GAA GAA ATT GAA TCC AGT GCT CAG TAC AAA Arg Ala Asn Gln Leu Ala Glu Glu Ile Glu Ser Ser Ala Gln Tyr Lys 260 265 270 | 871 |
| | GCT CGT GTC GCC CTT GAG AAT GAT GAC CGG AGT GAG GAA GAA AAA TAC Ala Arg Val Ala Leu Glu Asn Asp Asp Arg Ser Glu Glu Glu Lys Tyr 275 280 285 290 | 919 |
| | ACA GCA GTC CAG AGA AAC TGC AGT GAC CGG GAG GGG CAT GGC CCC AAC Thr Ala Val Gln Arg Asn Cys Ser Asp Arg Glu Gly His Gly Pro Asn 295 300 305 | 967 |
| | ACT AGG GAC AAT AAA TAT ATT CCT CCT GGA CAA AGA AAC AGA GAA GTC Thr Arg Asp Asn Lys Tyr Ile Pro Pro Gly Gln Arg Asn Arg Glu Val 310 315 320 | 1015 |
| | CTA TCC TGG GGA AGT GGG AGA CAG AGC TCA CCA CGG ATG GGC CAG CCT Leu Ser Trp Gly Ser Gly Arg Gln Ser Ser Pro Arg Met Gly Gln Pro 325 330 335 | 1063 |
| | GGG CCA GGC TCC ATG CCG TCA AGA GCT GCT TCT CAC ACT TCA GAT TTC Gly Pro Gly Ser Met Pro Ser Arg Ala Ala Ser His Thr Ser Asp Phe 340 345 350 | 1111 |
| | AAC CCG AAC GCT GGC TCA GAC CAA AGA GTA GTT AAT GGA GGT GTT CCC Asn Pro Asn Ala Gly Ser Asp Gln Arg Val Val Asn Gly Gly Val Pro 355 360 365 370 | 1159 |
| | TGG CCA TCG CCT TGC CCA TCT CAT TCC TCT CGC CCA CCT TCT CGC TAC Trp Pro Ser Pro Cys Pro Ser His Ser Ser Arg Pro Pro Ser Arg Tyr 375 380 385 | 1207 |
| | CAG TCA GGT CCC AAC TCT CTT CCA CCT CGG GCA GCC ACC CAT ACA CGG Gln Ser Gly Pro Asn Ser Leu Pro Pro Arg Ala Ala Thr His Thr Arg 390 395 400 | 1255 |
| | CCG CCC TCC AGG CCC CCC TCG AGG CCA TCC AGA CCC CCG TCT CAC CCC Pro Pro Ser Arg Pro Pro Ser Arg Pro Ser Arg Pro Pro Ser His Pro 405 410 415 | 1303 |
| | TCT GCT CAT GGT TCT CCA GCT CCT GTC TCT ACT ATG CCT AAA CGC ATG Ser Ala His Gly Ser Pro Ala Pro Val Ser Thr Met Pro Lys Arg Met 420 425 430 | 1351 |
| | TCT TCA GAA GGA CCC CCA AGG ATG TCT CCA AAG GCA CAG CGC CAC CCT Ser Ser Glu Gly Pro Pro Arg Met Ser Pro Lys Ala Gln Arg His Pro 435 440 445 450 | 1399 |
| | CGG AAT CAC AGA GTC TCT GCT GGG AGA GGC TCC ATG TCT AGT GGC CTA Arg Asn His Arg Val Ser Ala Gly Arg Gly Ser Met Ser Ser Gly Leu 455 460 465 | 1447 |
| | GAA TTT GTA TCC CAC AAT CCC CCA AGT GAA GCA GCT GCT CCT CCA GTG Glu Phe Val Ser His Asn Pro Pro Ser Glu Ala Ala Ala Pro Pro Val 470 475 480 | 1495 |
| | GCA AGG ACC AGT CCT GCA GGG GGA ACG TGG TCC TCA GTG GTC AGT GGG Ala Arg Thr Ser Pro Ala Gly Gly Thr Trp Ser Ser Val Val Ser Gly 485 490 495 | 1543 |

revised.SUB.SEQUENCE.LST.0120

| | |
|---|------|
| GTT CCA AGG TTA TCT CCC AAA ACT CAC AGA CCC AGG TCT CCC AGG CAG Val Pro Arg Leu Ser Pro Lys Thr His Arg Pro Arg Ser Pro Arg Gln 500 505 510 | 1591 |
| AGC AGC ATT GGA AAC TCT CCC AGC GGG CCT GTG CTT GCT TCT CCC CAA Ser Ser Ile Gly Asn Ser Pro Ser Gly Pro Val Leu Ala Ser Pro Gln 515 520 525 530 | 1639 |
| GCT GGC ATC ATC CCT GCA GAA GCC GTT TCC ATG CCT GTT CCC GCC GCA Ala Gly Ile Ile Pro Ala Glu Ala Val Ser Met Pro Val Pro Ala Ala 535 540 545 | 1687 |
| TCT CCG ACT CCT GCC AGC CCT GCA TCC AAC AGA GCA CTG ACC CCA TCT Ser Pro Thr Pro Ala Ser Pro Ala Ser Asn Arg Ala Leu Thr Pro Ser 550 555 560 | 1735 |
| ATT GAG GCA AAA GAT TCC AGG CTT CAA GAT CAG AGG CAG AAC TCT CCT Ile Glu Ala Lys Asp Ser Arg Leu Gln Asp Gln Arg Gln Asn Ser Pro 565 570 575 | 1783 |
| GCA GGG AGT AAA GAA AAT GTT AAA GCA AGT GAA ACA TCA CCT AGC TTT Ala Gly Ser Lys Glu Asn Val Lys Ala Ser Glu Thr Ser Pro Ser Phe 580 585 590 | 1831 |
| TCA AAA GCT GAC AAC AAA GGT ATG TCA CCA GTT GTT TCT GAA CAC AGA Ser Lys Ala Asp Asn Lys Gly Met Ser Pro Val Val Ser Glu His Arg 595 600 605 610 | 1879 |
| AAA CAG ATT GAT GAC TTA AAG AAG TTT AAG AAT GAT TTT AGG TTA CAG Lys Gln Ile Asp Asp Leu Lys Lys Phe Lys Asn Asp Phe Arg Leu Gln 615 620 625 | 1927 |
| CCA AGC TCT ACA TCT GAA TCT ATG GAT CAA CTA CTA AGC AAA AAT AGA Pro Ser Ser Thr Ser Glu Ser Met Asp Gln Leu Leu Ser Lys Asn Arg 630 635 640 | 1975 |
| GAA GGA GAA AAG TCA CGA GAT TTG ATT AAA GAT AAA ACG GAA GCA AGT Glu Gly Glu Lys Ser Arg Asp Leu Ile Lys Asp Lys Thr Glu Ala Ser 645 650 655 | 2023 |
| GCT AAG GAT AGT TTC ATT GAC AGC AGC AGC AGC AGC AGC AAC TGT ACC Ala Lys Asp Ser Phe Ile Asp Ser Ser Ser Ser Ser Ser Asn Cys Thr 660 665 670 | 2071 |
| AGT GGC AGC AGC AAG ACC AAC AGC CCT AGC ATC TCC CCT TCC ATG CTT Ser Gly Ser Ser Lys Thr Asn Ser Pro Ser Ile Ser Pro Ser Met Leu 675 680 685 690 | 2119 |
| AGT AAT GCA GAG CAC AAG AGG GGG CCT GAG GTC ACA TCC CAA GGG GTG Ser Asn Ala Glu His Lys Arg Gly Pro Glu Val Thr Ser Gln Gly Val 695 700 705 | 2167 |
| CAG ACT TCC AGC CCA GCC TGC AAA CAA GAG AAG GAT GAC AGA GAA GAG Gln Thr Ser Ser Pro Ala Cys Lys Gln Glu Lys Asp Asp Arg Glu Glu 710 715 720 | 2215 |
| AAG AAA GAC ACA ACA GAG CAG GTT AGG AAA TCG ACA TTG AAT CCC AAT Lys Lys Asp Thr Thr Glu Gln Val Arg Lys Ser Thr Leu Asn Pro Asn 725 730 735 | 2263 |

revised.SUB.SEQUENCE.LST.0120

| | |
|---|------|
| GCA AAG GAG TTC AAC CCT CGT TCT TTC TCT CAG CCA AAG CCT TCT ACT Ala Lys Glu Phe Asn Pro Arg Ser Phe Ser Gln Pro Lys Pro Ser Thr 740 745 750 | 2311 |
| ACC CCA ACG TCA CCT CGG CCT CAA GCA CAA CCC AGC CCA TCT ATG GTG Thr Pro Thr Ser Pro Arg Pro Gln Ala Gln Pro Ser Pro Ser Met Val 755 760 765 770 | 2359 |
| GGT CAT CAG CAG CCA GCT CCA GTG TAC ACT CAG CCT GTG TGC TTC GCA Gly His Gln Gln Pro Ala Pro Val Tyr Thr Gln Pro Val Cys Phe Ala 775 780 785 | 2407 |
| CCC AAT ATG ATG TAT CCC GTC CCA GTG AGC CCG GGC GTA CAA CCT TTA Pro Asn Met Met Tyr Pro Val Pro Val Ser Pro Gly Val Gln Pro Leu 790 795 800 | 2455 |
| TAC CCA ATA CCT ATG ACG CCC ATG CCT GTG AAC CAA GCC AAG ACA TAT Tyr Pro Ile Pro Met Thr Pro Met Pro Val Asn Gln Ala Lys Thr Tyr 805 810 815 | 2503 |
| AGA GCA GGT AAA GTA CCA AAT ATG CCC CAA CAG CGA CAA GAC CAA CAT Arg Ala Gly Lys Val Pro Asn Met Pro Gln Gln Arg Gln Asp Gln His 820 825 830 | 2551 |
| CAT CAA AGC ACC ATG ATG CAC CCA GCC TCC GCG GCA GGG CCA CCC ATC His Gln Ser Thr Met Met His Pro Ala Ser Ala Ala Gly Pro Pro Ile 835 840 845 850 | 2599 |
| GTA GCC ACC CCG CCC GCT TAC TCC ACT CAG TAC GTT GCC TAC AGC CCT Val Ala Thr Pro Pro Ala Tyr Ser Thr Gln Tyr Val Ala Tyr Ser Pro 855 860 865 | 2647 |
| CAG CAG TTT CCC AAT CAG CCT TTG GTC CAG CAT GTG CCG CAT TAT CAG Gln Gln Phe Pro Asn Gln Pro Leu Val Gln His Val Pro His Tyr Gln 870 875 880 | 2695 |
| TCT CAG CAT CCT CAT GTG TAC AGT CCT GTC ATA CAA GGT AAT GCC AGG Ser Gln His Pro His Val Tyr Ser Pro Val Ile Gln Gly Asn Ala Arg 885 890 895 | 2743 |
| ATG ATG GCA CCA CCA GCA CAT GCT CAG CCT GGT TTA GTG TCT TCT TCA Met Met Ala Pro Pro Ala His Ala Gln Pro Gly Leu Val Ser Ser Ser 900 905 910 | 2791 |
| GCT GCT CAG TTC GGG GCT CAC GAG CAG ACG CAC GCC ATG TAT GCA TGT Ala Ala Gln Phe Gly Ala His Glu Gln Thr His Ala Met Tyr Ala Cys 915 920 925 930 | 2839 |
| CCC AAA TTA CCA TAC AAC AAG GAG ACA AGC CCT TCT TTC TAC TTT GCC Pro Lys Leu Pro Tyr Asn Lys Glu Thr Ser Pro Ser Phe Tyr Phe Ala 935 940 945 | 2887 |
| ATT TCC ACC GGC TCC CTC GCT CAG CAG TAT GCA CAT CCT AAT GCC GCC Ile Ser Thr Gly Ser Leu Ala Gln Gln Tyr Ala His Pro Asn Ala Ala 950 955 960 | 2935 |
| CTG CAT CCA CAT ACT CCC CAT CCT CAG CCT TCG GCC ACT CCC ACC GGA Leu His Pro His Thr Pro His Pro Gln Pro Ser Ala Thr Pro Thr Gly 965 970 975 | 2983 |
| CAG CAG CAA AGC CAG CAT GGT GGA AGT CAC CCT GCA CCC AGT CCT GTT | 3031 |

revised.SUB.SEQUENCE.LST.0120

| | | |
|---|-------------------------------------|------|
| Gln Gln Gln Ser Gln His Gly | Gly Ser His Pro Ala Pro Ser Pro Val | |
| 980 | 985 990 | |
| CAG CAC CAT CAG CAC CAG GCT GCC CAG GCT CTT CAT CTG GCC AGT CCA | | 3079 |
| Gln His His Gln His Gln Ala Ala Gln Ala Leu His Leu Ala Ser Pro | 1000 1005 1010 | |
| 995 | | |
| CAG CAG CAG TCG GCC ATT TAT CAT GCG GGG CTG GCA CCA ACA CCA CCT | | 3127 |
| Gln Gln Gln Ser Ala Ile Tyr His Ala Gly Leu Ala Pro Thr Pro | 1015 1020 1025 | |
| | | |
| TCC ATG ACA CCT GCC TCT AAT ACA CAG TCT CCA CAG AGC AGT TTC CCA | | 3175 |
| Ser Met Thr Pro Ala Ser Asn Thr Gln Ser Pro Gln Ser Ser Phe Pro | 1030 1035 1040 | |
| | | |
| GCA GCA CAA CAG ACA GTC TTC ACC ATC CAC CCT TCT CAT GTT CAG CCG | | 3223 |
| Ala Ala Gln Gln Thr Val Phe Thr Ile His Pro Ser His Val Gln Pro | 1045 1050 1055 | |
| | | |
| GCA TAC ACC ACC CCA CCC CAC ATG GCC CAC GTA CCT CAG GCT CAT GTA | | 3271 |
| Ala Tyr Thr Thr Pro Pro His Met Ala His Val Pro Gln Ala His Val | 1060 1065 1070 | |
| | | |
| CAG TCA GGA ATG GTT CCT TCT CAT CCA ACT GCC CAT GCG CCA ATG ATG | | 3319 |
| Gln Ser Gly Met Val Pro Ser His Pro Thr Ala His Ala Pro Met Met | 1075 1080 1085 1090 | |
| | | |
| CTA ATG ACG ACA CAG CCA CCC GGT CCC AAG GCC GCC CTC GCT CAA AGT | | 3367 |
| Leu Met Thr Thr Gln Pro Pro Gly Pro Lys Ala Ala Leu Ala Gln Ser | 1095 1100 1105 | |
| | | |
| GCA CTA CAG CCC ATT CCA GTT TCG ACA ACA GCG CAT TTC CCT TAT ATG | | 3415 |
| Ala Leu Gln Pro Ile Pro Val Ser Thr Thr Ala His Phe Pro Tyr Met | 1110 1115 1120 | |
| | | |
| ACG CAC CCT TCA GTA CAA GCC CAC CAC CAA CAG CAG TTG TAA | | 3457 |
| Thr His Pro Ser Val Gln Ala His His Gln Gln Gln Leu | 1125 1130 1135 | |
| | | |
| GGCTGCCTTG GAGGAACCGA AAGGCCAAAT CCCTTCTTCC CTTCTCTGCT TCTGCCAAC | | 3517 |
| GGAAGCACAG AAAACTAGAA CTTCATTGAT TTTGTTTTTT AAAAGATACA CTGATTTAAC | | 3577 |
| ATCTGATAGG AATGCTAACA GCTCACTTGC AGTGGAGGAT GTTTTGGACC GAGTAGAGGC | | 3637 |
| ATGTAGGGAC TTGTGGCTGT TCCATAATTC CATGTGCTGT TGCAGGGTCC TGCAAGTACC | | 3697 |
| CAGCTCTGCT TGCTGAAACT GGAAGTTATT TATTTTTTAA TGGCCCTTGA GAGTCATGAA | | 3757 |
| CACATCAGCT AGCAACAGAA GTAACAAGAG TGATTCTTGC T | | 3798 |

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1135 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Ser Leu Lys Pro₅ Gln Pro Gln Pro₁₀ Ala Pro Ala Thr Gly Arg₁₅
 Lys Pro Gly Gly₂₀ Gly Leu Leu Ser₂₅ Ser Pro Gly Ala Ala Pro Ala Ser₃₀
 Ala Ala Val₃₅ Thr Ser Ala Ser Val₄₀ Val Pro Ala Pro₄₅ Ala Ala Pro Val₅₀
 Ala Ser Ser Ser Ala Ala Ala₅₅ Gly Gly Gly Arg Pro₆₀ Gly Leu Gly Arg₆₅
 Gly Arg Asn Ser Ser Lys₇₀ Gly Leu Pro Gln₇₅ Thr Ile Ser Phe Asp₈₀
 Gly Ile Tyr Ala Asn₈₅ Val Arg Met Val His₉₀ Ile Leu Thr Ser Val₉₅ Val₁₀₀
 Gly Ser Lys Cys₁₀₀ Glu Val Gln Val Lys₁₀₅ Asn Gly Gly Ile Tyr₁₁₀ Glu Gly₁₁₅
 Val Phe Lys Thr Tyr Ser Pro Lys₁₂₀ Cys Asp Leu Val Leu₁₂₅ Asp Ala Ala₁₃₀
 His Glu Lys Ser Thr Glu₁₃₅ Ser Ser Gly Pro Lys₁₄₀ Arg Glu Glu Ile₁₄₅
 Met Glu Ser Val Leu Phe₁₅₀ Lys Cys Ser Asp Phe₁₅₅ Val Val Val Gln Phe₁₆₀
 Lys Asp Thr Asp Ser₁₆₅ Ser Tyr Ala Arg Arg₁₇₀ Asp Ala Phe Thr Asp₁₇₅ Ser₁₈₀
 Ala Leu Ser Ala₁₈₀ Lys Val Asn Gly Glu₁₈₅ His Lys Glu Lys Asp₁₉₀ Leu Glu₁₉₅
 Pro Trp Asp₁₉₅ Ala Gly Glu Leu Thr₂₀₀ Ala Ser Glu Glu Leu₂₀₅ Glu Leu Glu₂₁₀
 Asn Asp Val Ser Asn Gly Trp₂₁₅ Asp Pro Asn Asp Met₂₂₀ Phe Arg Tyr Asn₂₂₅
 Glu Glu Asn Tyr Gly Val₂₃₀ Val Ser Thr Tyr Asp₂₃₅ Ser Ser Leu Ser₂₄₀ Ser₂₄₅
 Tyr Thr Val Pro Leu₂₄₅ Glu Arg Asp Asn Ser₂₅₀ Glu Glu Phe Leu Lys₂₅₅ Arg₂₆₀
 Glu Ala Arg Ala₂₆₀ Asn Gln Leu Ala Glu₂₆₅ Glu Ile Glu Ser Ser₂₇₀ Ala Gln₂₇₅
 Tyr Lys Ala Arg Val Ala Leu Glu₂₈₀ Asn Asp Asp Arg Ser₂₈₅ Glu Glu Glu₂₉₀
 Lys Tyr Thr Ala Val Gln₂₉₅ Arg Asn Cys Ser Asp Arg₃₀₀ Glu Gly His Gly₃₀₅
 Pro Asn Thr Arg Asp Asn₃₁₀ Lys Tyr Ile Pro Pro₃₁₅ Gly Gln Arg Asn Arg₃₂₀
 Glu Val Leu Ser Trp Gly Ser Gly Arg Gln Ser Ser Pro Arg Met Gly

325 330 335

Page 19

revised.SUB.SEQUENCE.LST.0120

Ala Ser Ala Lys Asp Ser Phe Ile Asp Ser Ser Ser Ser Ser Ser Asn
660 665 670

Cys Thr Ser Gly Ser Ser Lys Thr Asn Ser Pro Ser Ile Ser Pro Ser
675 680 685

Met Leu Ser Asn Ala Glu His Lys Arg Gly Pro Glu Val Thr Ser Gln
690 695 700

Gly Val Gln Thr Ser Ser Pro Ala Cys Lys Gln Glu Lys Asp Asp Arg
705 710 715 720

Glu Glu Lys Lys Asp Thr Thr Glu Gln Val Arg Lys Ser Thr Leu Asn
725 730 735

Pro Asn Ala Lys Glu Phe Asn Pro Arg Ser Phe Ser Gln Pro Lys Pro
740 745 750

Ser Thr Thr Pro Thr Ser Pro Arg Pro Gln Ala Gln Pro Ser Pro Ser
755 760 765

Met Val Gly His Gln Gln Pro Ala Pro Val Tyr Thr Gln Pro Val Cys
770 775 780

Phe Ala Pro Asn Met Met Tyr Pro Val Pro Val Ser Pro Gly Val Gln
785 790 795 800

Pro Leu Tyr Pro Ile Pro Met Thr Pro Met Pro Val Asn Gln Ala Lys
805 810 815

Thr Tyr Arg Ala Gly Lys Val Pro Asn Met Pro Gln Gln Arg Gln Asp
820 825 830

Gln His His Gln Ser Thr Met Met His Pro Ala Ser Ala Ala Gly Pro
835 840 845

Pro Ile Val Ala Thr Pro Pro Ala Tyr Ser Thr Gln Tyr Val Ala Tyr
850 855 860

Ser Pro Gln Gln Phe Pro Asn Gln Pro Leu Val Gln His Val Pro His
865 870 875 880

Tyr Gln Ser Gln His Pro His Val Tyr Ser Pro Val Ile Gln Gly Asn
885 890 895

Ala Arg Met Met Ala Pro Pro Ala His Ala Gln Pro Gly Leu Val Ser
900 905 910

Ser Ser Ala Ala Gln Phe Gly Ala His Glu Gln Thr His Ala Met Tyr
915 920 925

Ala Cys Pro Lys Leu Pro Tyr Asn Lys Glu Thr Ser Pro Ser Phe Tyr
930 935 940

Phe Ala Ile Ser Thr Gly Ser Leu Ala Gln Gln Tyr Ala His Pro Asn
945 950 955 960

Ala Ala Leu His Pro His Thr Pro His Pro Gln Pro Ser Ala Thr Pro
965 970 975

Thr Gly Gln Gln Gln Ser Gln His Gly Gly Ser His Pro Ala Pro Ser
980 985 990

revised.SUB.SEQUENCE.LST.0120

Pro Val Gln His His Gln His Gln Ala Ala Gln Ala Leu His Leu Ala
 995 1000 1005
 Ser Pro Gln Gln Gln Ser Ala Ile Tyr His Ala Gly Leu Ala Pro Thr
 1010 1015 1020
 Pro Pro Ser Met Thr Pro Ala Ser Asn Thr Gln Ser Pro Gln Ser Ser
 1025 1030 1035 1040
 Phe Pro Ala Ala Gln Gln Thr Val Phe Thr Ile His Pro Ser His Val
 1045 1050 1055
 Gln Pro Ala Tyr Thr Thr Pro Pro His Met Ala His Val Pro Gln Ala
 1060 1065 1070
 His Val Gln Ser Gly Met Val Pro Ser His Pro Thr Ala His Ala Pro
 1075 1080 1085
 Met Met Leu Met Thr Thr Gln Pro Pro Gly Pro Lys Ala Ala Leu Ala
 1090 1095 1100
 Gln Ser Ala Leu Gln Pro Ile Pro Val Ser Thr Thr Ala His Phe Pro
 1105 1110 1115 1120
 Tyr Met Thr His Pro Ser Val Gln Ala His His Gln Gln Gln Leu
 1125 1130 1135

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGGCCCCTCA CCATGTCG

18

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGGCTTGCG GACATTGG

18

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCGCGGCTGC CAATGTCC

18

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GTAACCGTTC GGCGCCCG

18

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGCTCCCGGC GGCTCCTT

18

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGCTGCTGCT GCTGGGGCTT CAG

23

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCGCCCCGCTC CTCACGTGT

19

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ACCCCCGAGA AAGCAACC

18

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCGTTGCCGT TGCTACCA

18

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TTCTCATGTG CGGCATCAAG

20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 346 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```

Val Tyr Gly Pro Leu Thr Met Ser Leu Lys Pro Gln Gln Gln Gln Gln
1      5      10      15
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
20     25     30
Gln Pro Pro Pro Ala Ala Ala Asn Val Arg Lys Pro Gly Gly Ser Gly
35     40     45
Leu Leu Ala Ser Pro Ala Ala Ala Pro Ser Pro Ser Ser Ser Val
50     55     60
Ser Ser Ser Ser Ala Thr Ala Pro Ser Ser Val Val Ala Ala Thr Ser
65     70     75     80
Gly Gly Gly Arg Pro Gly Leu Gly Arg Gly Arg Asn Ser Asn Lys Gly
85     90     95
Leu Pro Gln Ser Thr Ile Ser Phe Asp Gly Ile Tyr Ala Asn Met Arg
100    105    110
Met Val His Ile Leu Thr Ser Val Val Gly Ser Lys Cys Glu Val Gln
115    120    125
Val Lys Asn Gly Gly Ile Tyr Glu Gly Val Phe Lys Thr Tyr Ser Pro
130    135    140
Lys Cys Asp Leu Val Leu Asp Ala Ala His Glu Lys Ser Thr Glu Ser
145    150    155    160
Ser Ser Gly Pro Lys Arg Glu Glu Ile Met Glu Ser Ile Leu Phe Lys
165    170    175
Cys Ser Asp Phe Val Val Val Gln Phe Lys Asp Met Asp Ser Ser Tyr
180    185    190

```


revised.SUB.SEQUENCE.LST.0120

Ala Lys Arg Asp Ala Phe Thr Asp Ser Ala Ile Ser Ala Lys Val Asn
195 200 205

Gly Glu His Lys Glu Lys Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu
210 215 220

Thr Ala Asn Glu Glu Leu Glu Ala Leu Glu Asn Asp Val Ser Asn Gly
225 230 235 240

Trp Asp Pro Asn Asp Met Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val
245 250 255

Val Ser Thr Tyr Asp Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu
260 265 270

Arg Asp Asn Ser Glu Glu Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln
275 280 285

Leu Ala Glu Glu Ile Glu Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala
290 295 300

Leu Glu Asn Asp Asp Arg Ser Glu Glu Glu Lys Tyr Thr Ala Val Gln
305 310 315 320

Arg Asn Ser Ser Glu Arg Glu Gly His Ser Ile Asn Thr Arg Glu Asn
325 330 335

Lys Tyr Ile Pro Pro Gly Gln Arg Asn Arg
340 345

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 326 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

His Glu Gly Pro Leu Thr Met Ser Leu Lys Pro Gln Pro Gln Pro Pro
1 5 10 15

Ala Pro Ala Thr Gly Arg Lys Pro Gly Gly Gly Leu Leu Ser Ser Pro
20 25 30

Gly Ala Ala Pro Ala Ser Ala Ala Val Thr Ser Ala Ser Val Val Pro
35 40 45

Ala Pro Ala Ala Pro Val Ala Ser Ser Ser Ala Ala Ala Gly Gly Gly
50 55 60

Arg Pro Gly Leu Gly Arg Gly Arg Asn Ser Ser Lys Gly Leu Pro Gln
65 70 75 80

Pro Thr Ile Ser Phe Asp Gly Ile Tyr Ala Asn Val Arg Met Val His
85 90 95

revised.SUB.SEQUENCE.LST.0120

Ile Leu Thr Ser Val Val Gly Ser Lys Cys Glu Val Gln Val Lys Asn
100 105 110
Gly Gly Ile Tyr Glu Gly Val Phe Lys Thr Tyr Ser Pro Lys Cys Asp
115 120 125
Leu Val Leu Asp Ala Ala His Glu Lys Ser Thr Glu Ser Ser Ser Gly
130 135 140
Pro Lys Arg Glu Glu Ile Met Glu Ser Val Leu Phe Lys Cys Ser Asp
145 150 155 160
Phe Val Val Val Gln Phe Lys Asp Thr Asp Ser Ser Tyr Ala Arg Arg
165 170 175
Asp Ala Phe Thr Asp Ser Ala Leu Ser Ala Lys Val Asn Gly Glu His
180 185 190
Lys Glu Lys Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu Thr Ala Ser
195 200 205
Glu Glu Leu Glu Leu Glu Asn Asp Val Ser Asn Gly Trp Asp Pro Asn
210 215 220
Asp Met Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val Val Ser Thr Tyr
225 230 235 240
Asp Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Arg Asp Asn Ser
245 250 255
Glu Glu Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln Leu Ala Glu Glu
260 265 270
Ile Glu Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala Leu Glu Asn Asp
275 280 285
Asp Arg Ser Glu Glu Glu Lys Tyr Thr Ala Val Gln Arg Asn Cys Ser
290 295 300
Asp Arg Glu Gly His Gly Pro Asn Thr Arg Asp Asn Lys Tyr Ile Pro
305 310 315 320
Pro Gly Gln Arg Asn Arg
325

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 275 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Leu Ala Pro Gln Pro Pro Pro Pro Gln Gln His Gln Glu Arg Pro Gly
1 5 10 15

revised.SUB.SEQUENCE.LST.0120

Ala Ala Ala Ile Gly Ser Ala Arg Gly Gln Ser Thr Gly Lys Gly Pro
20 25 30
Pro Gln Ser Pro Val Phe Glu Gly Val Tyr Asn Asn Ser Arg Met Leu
35 40 45
His Phe Leu Thr Ala Val Val Gly Ser Thr Cys Asp Val Lys Val Lys
50 55 60
Asn Gly Thr Thr Tyr Glu Gly Ile Phe Lys Thr Leu Ser Ser Lys Phe
65 70 75 80
Glu Leu Ala Val Asp Ala Val His Arg Lys Ala Ser Glu Pro Ala Gly
85 90 95
Gly Pro Arg Arg Glu Asp Ile Val Asp Thr Met Val Phe Lys Pro Ser
100 105 110
Asp Val Met Leu Val His Phe Arg Asn Val Asp Phe Asn Tyr Ala Thr
115 120 125
Lys Asp Lys Phe Thr Asp Ser Ala Ile Ala Met Asn Ser Lys Val Asn
130 135 140
Gly Glu His Lys Glu Lys Val Leu Gln Arg Trp Glu Gly Gly Asp Ser
145 150 155 160
Asn Ser Asp Asp Tyr Asp Leu Glu Ser Asp Met Ser Asn Gly Trp Asp
165 170 175
Pro Asn Glu Met Phe Lys Phe Asn Glu Glu Asn Tyr Gly Val Lys Thr
180 185 190
Thr Tyr Asp Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Lys Asp
195 200 205
Asn Ser Glu Glu Phe Arg Gln Arg Glu Leu Arg Ala Ala Gln Leu Ala
210 215 220
Arg Glu Ile Glu Ser Ser Pro Gln Tyr Arg Leu Arg Ile Ala Met Glu
225 230 235 240
Asn Asp Asp Gly Arg Thr Glu Glu Glu Lys His Ser Ala Val Gln Arg
245 250 255
Gln Gly Ser Gly Arg Glu Ser Pro Ser Leu Ala Ser Arg Glu Gly Lys
260 265 270
Tyr Ile Pro
275

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1257 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 2..1255

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

| | |
|---|-----|
| G CAC GAG GGG CCG CTC ACC ATG TCG CTG AAG CCG CAG CCG CAG CCG | 46 |
| His Glu Gly Pro Leu Thr Met Ser Leu Lys Pro Gln Pro Gln Pro | |
| 1 5 10 15 | |
| CCC GCG CCC GCC ACT GGC CGC AAG CCC GGC GGC GGC CTG CTC TCG TCG | 94 |
| Pro Ala Pro Ala Thr Gly Arg Lys Pro Gly Gly Gly Leu Leu Ser Ser | |
| 20 25 30 | |
| CCC GGC GCC GCG CCG GCC TCG GCC GCG GTG ACC TCG GCT TCC GTG GTG | 142 |
| Pro Gly Ala Ala Pro Ala Ser Ala Ala Val Thr Ser Ala Ser Val Val | |
| 35 40 45 | |
| CCG GCC CCG GCC GCG CCG GTG GCG TCT TCC TCG GCG GCC GCG GGC GGC | 190 |
| Pro Ala Pro Ala Ala Pro Val Ala Ser Ser Ser Ala Ala Ala Gly Gly | |
| 50 55 60 | |
| GGG CGT CCC GGC CTG GGC AGA GGT CGG AAC AGT AGC AAA GGA CTG CCT | 238 |
| Gly Arg Pro Gly Leu Gly Arg Gly Arg Asn Ser Ser Lys Gly Leu Pro | |
| 65 70 75 | |
| CAG CCT ACG ATT TCT TTT GAT GGA ATC TAT GCA AAC GTG AGG ATG GTT | 286 |
| Gln Pro Thr Ile Ser Phe Asp Gly Ile Tyr Ala Asn Val Arg Met Val | |
| 80 85 90 95 | |
| CAT ATA CTT ACG TCA GTT GTT GGA TCG AAA TGT GAA GTA CAA GTG AAA | 334 |
| His Ile Leu Thr Ser Val Val Gly Ser Lys Cys Glu Val Gln Val Lys | |
| 100 105 110 | |
| AAC GGA GGC ATA TAT GAA GGA GTT TTT AAA ACA TAC AGT CCT AAG TGT | 382 |
| Asn Gly Gly Ile Tyr Glu Gly Val Phe Lys Thr Tyr Ser Pro Lys Cys | |
| 115 120 125 | |
| GAC TTG GTA CTT GAT GCT GCA CAT GAG AAA AGT ACA GAA TCC AGT TCG | 430 |
| Asp Leu Val Leu Asp Ala Ala His Glu Lys Ser Thr Glu Ser Ser Ser | |
| 130 135 140 | |
| GGG CCA AAA CGT GAA GAA ATA ATG GAG AGT GTT TTG TTC AAA TGC TCA | 478 |
| Gly Pro Lys Arg Glu Glu Ile Met Glu Ser Val Leu Phe Lys Cys Ser | |
| 145 150 155 | |
| GAC TTC GTT GTG GTA CAG TTT AAA GAT ACA GAC TCC AGT TAT GCA CGG | 526 |
| Asp Phe Val Val Val Gln Phe Lys Asp Thr Asp Ser Ser Tyr Ala Arg | |
| 160 165 170 175 | |
| AGA GAT GCT TTT ACT GAC TCT GCT CTC AGC GCA AAG GTG AAT GGT GAG | 574 |
| Arg Asp Ala Phe Thr Asp Ser Ala Leu Ser Ala Lys Val Asn Gly Glu | |
| 180 185 190 | |
| CAC AAG GAG AAG GAC CTG GAG CCC TGG GAT GCA GGG GAG CTC ACG GCC | 622 |
| His Lys Glu Lys Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu Thr Ala | |
| 195 200 205 | |
| AGC GAG GAG CTG GAG CTG GAG AAT GAT GTG TCT AAT GGA TGG GAC CCC | 670 |
| Ser Glu Glu Leu Glu Leu Glu Asn Asp Val Ser Asn Gly Trp Asp Pro | |
| 210 215 220 | |

revised.SUB.SEQUENCE.LST.0120

| | |
|---|------|
| AAT GAC ATG TTT CGA TAT AAT GAA GAG AAT TAT GGT GTG GTG TCC ACA Asn Asp Met Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val Val Ser Thr 225 230 235 | 718 |
| TAT GAT AGC AGT TTA TCT TCA TAT ACG GTT CCT TTA GAA AGG GAC AAC Tyr Asp Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Arg Asp Asn 240 245 250 255 | 766 |
| TCA GAA GAA TTT CTT AAA CGG GAG GCA AGG GCA AAC CAG TTA GCA GAA Ser Glu Glu Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln Leu Ala Glu 260 265 270 | 814 |
| GAA ATT GAA TCC AGT GCT CAG TAC AAA GCT CGT GTC GCC CTT GAG AAT Glu Ile Glu Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala Leu Glu Asn 275 280 285 | 862 |
| GAT GAC CGG AGT GAG GAA GAA AAA TAC ACA GCA GTC CAG AGA AAC TGC Asp Asp Arg Ser Glu Glu Glu Lys Tyr Thr Ala Val Gln Arg Asn Cys 290 295 300 | 910 |
| AGT GAC CGG GAG GGG CAT GGC CCC AAC ACT AGG GAC AAT AAA TAT ATT Ser Asp Arg Glu Gly His Gly Pro Asn Thr Arg Asp Asn Lys Tyr Ile 305 310 315 | 958 |
| CCT CCT GGA CAA AGA AAC AGA GAA GTC CTA TCC TGG GGA AGT GGG AGA Pro Pro Gly Gln Arg Asn Arg Glu Val Leu Ser Trp Gly Ser Gly Arg 320 325 330 335 | 1006 |
| CAG AGC TCA CCA CGG ATG GGC CAG CCT GGG CCA GGC TCC ATG CCG TCA Gln Ser Ser Pro Arg Met Gly Gln Pro Gly Pro Gly Ser Met Pro Ser 340 345 350 | 1054 |
| AGA GCT GCT TCT CAC ACT TCA GAT TTC AAC CCG AAC GCT GGC TCA GAC Arg Ala Ala Ser His Thr Ser Asp Phe Asn Pro Asn Ala Gly Ser Asp 355 360 365 | 1102 |
| CAA AGA GTA GTT AAT GGA GGT GTT CCC TGG CCA TCG CCT TGC CCA TCT Gln Arg Val Val Asn Gly Gly Val Pro Trp Pro Ser Pro Cys Pro Ser 370 375 380 | 1150 |
| CCT TCC TCT CGC CCA CCT TCT CGC TAC CAG TCA GGT CCC AAC TCT CTT Pro Ser Ser Arg Pro Pro Ser Arg Tyr Gln Ser Gly Pro Asn Ser Leu 385 390 395 | 1198 |
| CCA CCT CGG GCA GCC ACC CCT ACA CGG CCT CGT GCC GAA TTC CTG CAG Pro Pro Arg Ala Ala Thr Pro Thr Arg Pro Arg Ala Glu Phe Leu Gln 400 405 410 415 | 1246 |
| CCC GGG GAT CC Pro Gly Asp | 1257 |